

Kirjallisuutta

Uutta viron kielen tutkimusta

ÜLLE VIKS *Klassifikatoorne morfoloogia: verb.* Valgus, Tallinn 1980. 68 s.

MATI ERELT *Adjektiivatribuut eesti keeles.* Valgus, Tallinn 1980. 56 s.

Viron tiedeakatemian kielen ja kirjallisuuden instituutti on viime vuosina julkaissut vihkosarjana useita kiintoisia tutkielmia, joissa käsitellään lähinnä nykyviron yleiskielen kysymyksiä; eräissä näkökulma on laajempikin: yleiskielitieteellinen. Vihkoja on tähän mennessä ilmestynyt tietääkseni jo toistakymmentä. Tarkastelen tässä niistä kahta.

Ülle Viksin tutkimuskohtena on morfoloogia, jota lähestytään tietokonekielipöpin näkökulmasta, mikä ihmiskielioppiin nähden merkitsee suurempia formaalisuuden ja eksplisiittisyyden vaatimuksia. Tutkielmassaan tekijä soveltaa kehittämänsä morfologisen luokituksen mallia viron verbimorfologiaan. Hän on julkaissut

aiemmin tässä samassa sarjassa mallinsa teoriapuolta käsittelevän kirjoituksen *Klassifikatoorse morfoloogia põhimõtted* (lyh. KMP, Tallinn 1977), josta poimin aluksi pari ydinkohtaa taustaksi. Morfologisessa mallissaan Viks pyrkii yhdistämään systeemin, struktuurin ja funktion käsittelyn: hänen morfoloigiansa on paradigmaattis-syntagmaattista. Morfologista struktuuria kuvataan laatimalla sanojen morfologinen luokitus kriteereinä nimenomaan sellaiset tunnuksat (eli piirteet, jotka ovat ominaisia sanan taivutukselle ja erottavat sen samalla tiettyjen muiden sanojen taivutuksesta), jotka ovat systeemin toiminnan kannalta tärkeitä, esim. syntetisoitaessa sanavartalo konkreettisiksi paradigmaamuodoiksi. Morfologisen systeemin toiminnan kuvauksessa tarvitaan taulukon muotoon laaditun luokituksen lisäksi sääntöjä, esim. päätteiden sijoitussääntöjä ja astevaihtelusääntöjä sekä yleensä sanasto. Jotta morfologinen kuvaus kokonaisuutena olisi mahdollisimman yksinkertainen, ekonominen ja yleistävä, on löydettävä järkevää tasapaino informaation jakamisessa luokituksen, sääntöjen ja poikkeusluetteloiden kesken. (Mts. 23—29.)

Siirryn nyt itse morfologisen luokituksen tarkasteluun. Luokitusmetodin perusteet tekijä on omaksunut tekoälyn tutkimuksen piiriin kuuluvasta hahmontunnistusteoriasta. Sen mukaan voi morfologiselle luokittelulle asettaa seuraavat periaatteet ja tavoitteet: 1) morfologisten tunnusten selvittäminen, 2) sanojen alustava luokitus, ensin kunkin löydetyn tunnukseen mukaan erikseen ja sitten kaikkien mukaan samalla kertaa, 3) luokituksen optimointi niin, että jäljelle jäävät olennaiset tunnuksat ja tyypilliset luokat, 4) yksityisen sanan luokan identifiointi: miten sanan oma luokka valmiissa luokitus-systeemissä voidaan tunnistaa, 5) luokituksen arviointi mm. adekvaattiutta, yleistävyyttä, ekonomisuutta ja systeemin toimintanopeutta silmällä pitäen. Tekijän mielestä hahmontunnistusteoria sopii erittäin hyvin viron morfoloigian kuvaukseen, koska sen keskeiset tavoitteet, klassifiointi ja identifiointi, muistuttavat paljon viron perinnäisen morfologisen kuvauksen päätavoitteita: sanojen jakoa taivutustyy-

peiksi ja tyyppisanan määrittämistä.

Viron morfologisessa luokituksessa tarvittavia tunnuksia on kahta tyyppiä. Ne koskevat toisaalta vartalossa ja toisaalta päätteissä eli kieliopillisissa formatiiveissa esiintyvää morfologista vaihtelua. Morfologisessa luokituksessa ei oteta huomioon teksti- eli äännetasolla usein esiintyvää morfofonologista vaihtelua, jonka ilmentymiksi tässä katsotaan vartalon ja formatiivin rajaiset äännesekvenssien assimilaatio- ja yksinkertaistumistapaukset ym. Tunnuksista ensimmäinen koskee formatiiveja. Esim. *da*-infinitiivin formatiivilla on kolme erilaista morfologista varianttia, joiden mukaan verbit luokituvat kolmia: tyyppit *elada*, *hakata*, *t'uua* (merkintä 3. kestoastetta). Sellainen tapaus kuin *'anda* kuuluu samaan tyyppiin kuin *elada*, koska sen *a*-loppu katsotaan morfofonologiseksi variantiksi; morfologisella tasolla on kyse *da*-loppuisesta muodosta (ns. morfofonologisen distributiosäännön avulla saadaan *änd+da* → *'anda*). Seuraavat kaksi tunnusta liittyvät astevaihteluun. Luokituksessa on tekijän mukaan otettava huomioon, esiintyykö sanassa astevaihtelua vai ei, ja jos esiintyy, niin kesto- vai lauvaihtelua, koska tätä ei virossa voi läheskään aina ennustaa sanan äännerakenteen nojalla, vrt. esim. *tõdema* : *tõdeb* — *põdema* : *p'õeb*. Astevaihtelun esiintyminen on siis pitkälle leksikaalistunut ilmiö. Sen sijaan vahvan ja heikon asteen sijoittuminen paradigmaan on jo puhtaasti morfologinen tunnus, vrt. esim. *'õppima* : *'õppida* : *õpin* — *h'akkama* : *hakata* : *h'akkan*. *'õppima*- ja *h'akkama*-verbi kuuluvat tämän tunnuksen perusteella eri tyyppeihin. Samantapainen on viimeinen tunnus: vartalon loppuäänteitä koskeva ns. vartalovaihtelu eli lemman- ja taivutusvartalon sijoittuminen paradigmaan (lemmavartalolla tarkoitetaan vartalon perusmuotoa: verbeillä *ma*-infinitiiviä), vrt. esim. *s'eis/ma* : *s'eis/ta* : *seisa/b* — *kõnele/ma* : *kõnel/da* : *kõnele/b*. Astevaihtelutapauksissa voi vahvasta asteesta muodostaa säännön avulla heikon. Vartalovaihtelutapauksissa taas ei voi säännöin johtaa esim. lemmavartalosta taivutusvartaloa varsinkaan nomaineilla, vrt. esim. *ehe* : *eheda* 'puhdas' — *ehe* : *'ehte* 'koru' — *ehe* : *'ehtme* 'selän keskiosa'. Tämän vuoksi konstruoidaan kullekin sa-

nalle semmoinen pohjamuoto, joka sisältää suluissa kaikki vartalovaihtelun alaiset äänne-elementit niin, että ensimmäisten sulkujen sisältö kuuluu lemman-, toisten taivutusvartaloon; lisäksi se tulee vahvan asteeseen (jos sana on astevaihtelun alainen): *KÕNEL(E)()*, *EHE()*(*DA*), *'EHTE* (tässä ei ole vartalovaihtelua), *'EHT(E)*(*ME*). Pohjamuoto toimii leksikkomuotona ja syötösmuotona synteesiprosessissa.

Viksin aineistona on Öigekeelsussõnaraamat (1976) -teoksen 8 073 verbien perusparadigmat: 46 muotoa (finiittiset ja infiniittiset muodot paitsi liittomuodot) kustakin verbistä. Koko verbistö luokitetaan erikseen kunkin morfologisen tunnuksen suhteen mahdollisimman optimaalisesti taivutustyypeiksi, jotka numeroidaan. Näin kukin verbi saa neljä tyyppinumeroa. Lopuksi osaluokitukset koetaan yhdistelmäluokitukseksi, jossa samaan tyyppiin tulevat verbit, joilla on sama numerokombinaatio. Näin tähän yhdistelmäluokitukseen (taulukko sivulla 36) tulee 17 numeroitua taivutustyyppiä. Verbin tyyppinumero liitetään sanastoon pohjamuodon ohkeen lisäinformaatioksi. Voinee sanoa, että tyyppinumero osoittaa päähkinänkuoressa verbin aseman morfologisessa struktuurissa ja ohjaa sen käytäytymistä systeemin toiminnassa.

Tässä on mahdollista tarkastella lähempää vain lopullista, yhdistelmäluokitusta. Se jakaantuu kolmeksi konjugaatioksi, joista I:een kuuluu 12 taivutustyyppiä (joissa yhteensä 7 978 verbiä 8 073:sta), II:een vain yksi tyyppi (77 verbiä) sekä III:een neljä tyyppiä (yhteensä 18 verbiä). III konjugaation verbit ovat yksitavuisia vokaalivartaloja. Ne jakaantuvat tyypeiksi formatiivien ja sen mukaan, miten vahva ja heikko aste sijoittuvat paradigmaan, esim. *käima*-tyyppi: *käia* : *käiakse* : *käisin*, vrt. *s'aama*-tyyppi: *s'aada* : *s'aadakse* : *s'ain*. II konjugaation ainoa tyyppi käsittää keskeiset verbit *tulema*, *olema*, *minema*, *panema*, *surema* ja joukon harvinaisia verbejä, esim. *tantsikelema*. Näissä esiintyy kestoastevaihtelu, samoin vartalovaihtelu, esim. *tule/ma* : *t'ull/a* : *tull/akse* (perinnäiset kieliopit eivät ota huomioon näiden eivätkä myöskään III konjugaation verbien astevaihtelua). Voi mainita, että esim.

Kirjallisuutta

Valgman—Remmelin Eesti keele grammatika -teoksessa (lyh. EKG, Tallinn 1968) *tulema*-tyyppiset verbit luetaan *kõnelema*-tyyppiin (joka Viksillä kuuluu I konjugaatioon): ne eroavat tästä (astevaihtelun lisäksi) lähinnä imperfektimuotojen formaatiivien osalta: *tulin* mutta *kõnelesin*, ja lisäksi esim. *da*-infinitiiviltään: *t'ull/a* mutta *kõnel/da*. Eroa voisi pientää analysoimalla *t'ulla*-tyypin *da*-infinitiivin näin: *t'ul/la* (*t'ul+da* → *t'ulla*; tällainen morfonologinen sääntö ei liene lingvistisesti vähemmän perusteltavissa kuin Viksin käyttämä *t'ul+a* → *t'ulla*). Silti täytyy ainakin Viksin lähtökohdista asiaa tarkastellen pitää nämä tyypit erillään; yhdistäminen näet tuotaisi tuntuvasti poikkeuksia, ja sitä paitsi *tulema*-tyypin status lienee aika vahva, koska tyypin muutamata verbit ovat keskeistä sanastoa. I konjugaatioon suurin ja samalla koko verbintaivutuksen produktiivisin tyyppi on *elama*-tyyppi (5 044 verbiä). Siinä taivutus on yksinkertaisinta: ei ole aste eikä vartalovaihtelua, ja formatiivit esiintyvät sanoisinko tuntomerkittömässä asussaan. *õppima*-tyypin (1 929) erottaa tästä kestovaihtelun esiintyminen, *h'akkama*-tyyppi (449) erottuu lisäksi formatiivien osalta. Mutkikkainta on taivutus *p'iinlema* (133) ja *õmblema*-tyypissä (58), joissa vaikuttaa lisäksi vartalovaihtelu. (Taivutuksen mutkikkuus siis kuvastuu tyyppien koossa.) Yhteensä kymmenen I konjugaatioon tyyppiä ryhmitetty pareiksi siten, että erottavana tunnuksena on astevaihtelun laji, esim. *õppima* — *s'ündima*, *s'aatma* — *t'ahtma* (viimeiseen kuuluu vain kolme verbiä).

Saatu luokitus on ilmeisen optimaalinen: taivutustyyppiä on semmoinen määrä (17, EKG:n neljässä konjugaatiossa yhteensä 14 tyyppiä, vrt. mts. 127—132), että sääntöjen ulottumattomiin jääviä poikkeusmuotoja on luetteloitava vain pariltakymmeneltä verbiltä yhteensä 52 (eniten *minema*-, *nägema*- ja *tegeme*-verbeistä). Tekijä toteaa, että III konjugaatioon pieniä tyyppiä ei voi yhdistellä, koska poikkeusmuotojen suhteellinen määrä tulisi niissä liian suureksi. Myöskään *tulema*-tyyppiä ei voine eliminoida, kuten edellä totesin. Mutta voisi harkita sitä, että poistettaisiin astevaihtelun esiintymisen ja lajin osoittaminen morfologisten tunnusten joukosta, koska kyseessä on paljon leksi-

kaalinen ilmiö (itse asiassa Viks ei mainitse tätä aiemmassa tutkielmassaan morfologisten tunnusten joukossa, vrt. KMP s. 36). Tunnuksen voisi liittää sanastoon tyyppinumeron oheen lisäinformaatioksi; vartalovaihtelun esiintyminen ja lajihän esitetään sanastossa, mikä tosin sen kohdalla onkin välttämätöntä. Tyyppien lukumäärä laskee tällöin 12:een, kun I konjugaatioon jo edellä mainitut tyypit (esim. *õppima* — *s'ündima*) lankeavat pareittain yhteen. Tämä ei tietenkään tuota uusia poikkeusmuotoja eikä mutkistane synteesin kulkua.

Ülle Viksin suorittama viron verbien luokitus ei poikkea perinnäisen kieliopin vastaavasta paljonkaan. Kriteerit ovatkin käytännössä enimmäkseen samoja: astevaihtelun esiintyminen, asteiden sijainti paradigmassa, mistä EKG käyttää termejä heikkenevä ja vahveneva astevaihtelu (lähtökohtana *da*-infinitiivi), ja vartalovaihtelu, jonka EKG osoittaa äänteellisin tunnuksin, esim. toteamalla, että *õmblema*-tyypissä *da*-infinitiivin vartalo on *l*-loppuinen. EKG ei käytä formatiivikriteeriä vaan osoittaa tyypeittäin formatiivien asut tyyppisanan teemamuotojen avulla.

Esitelty morfologinen luokitus on adekvaatti, koska tekijä on voinut käytännössä todeta, että se täyttää päätehtävänsä: sen avulla voi syntetisoida virheettömästi viron verbien yksityisiä taivutusmuotoja ja kokonaisia paradigmoja. Päinvastaisen prosessin suorittaminen eli tekstissä esiintyvien sanamuotojen analysointi leksikaalista ja kieliopillista merkitystä kannattaviksi yksiköiksi on paljon visaisempi ongelma etenkin viron kaltaisessa morfoloogiaan pitkälle ei-agglutinovassa ja (varsinkin nominintaivutuksessa) muotojen homonymiaa sisältävässä kielessä. Tämä vaatisi, kuten tekijä toteaa, mutkikasta morfonologista luokitusta, jollaisen yleistysarvo olisi kuitenkin pieni. Viksin työ on joka tapauksessa merkittävä alkuaskel viron morfologian automaattisessa käsittelyssä.

Mati Erelin syntaktis-semanttisen tutkielman aiheena ovat viron adjektiiviattribuuttirakenteet. Siinä käsitellään koordinaatiivisia ja subordinatiivisia attribuuttirakenteita ja niiden erottamiseen liittyviä ongelmia, samoin adjektiiviattribuutin

restrktiivistä eli klassifioivaa funktiota verrattuna predikatiivilauseessa tavattavan adjektiivin ei-restrktiiviseen eli kvalifioivaan funktioon. Keskeistä ja ehkä eniten uutta tarjoavaa julkaisussa on kumminkin attribuuttirakenteiden tutkiminen proposition näkökulmasta. Keskityn seuraavassa tämän näkökohdan tarkasteluun.

Attribuuttirakenteen rinnalla esiintyy usein predikatiivilause: *pikk poiss — Poiss on pikk*. Kaikkiin rakenteisiin tämä ei kuitenkaan päde: *täielik bürokraat — *(See) bürokraat on täielik, jooksev poiss — *Poiss on jooksev*. Attribuuttirakenteita ei siis ainakaan yleisesti voi pitää predikatiivilauseiden nominalisaatioina (kuten klassinen TG-kielioppi teki). Ereltin mukaan se abstraktio, jonka kannalta kaikkia attribuuttirakenteita (ja muitakin lauseen rakenteita) on tarkasteltava, on niiden proposition rakenne. Ensimmäisen esimerkin attribuuttirakenteella ja sitä vastaavalla predikatiivilauseella on syvyydessä yhteinen proposition rakenne (ei siis katsota, että edellinen johtuisi jälkimmäisestä); pinnalla niiden ero on kommunikatiivis-pragmaattinen. Muiden esimerkkien attribuuttirakenteilla on omat, tästä eroavat proposition rakenteensa. Attribuuttirakenteet voidaan proposition kannalta jakaa neljään tyyppiin.

Perustyyppin rakenteissa attribuutin voi käsittää ominaisuutta ilmaisevaksi argumentiksi, esim. *pikk poiss, punane lipp, kingituse üle rõõmus laps*. Perustyyppissä on attribuuttina primaari adjektiivi, muissa tyypeissä taas leksikaalis-semanttisesti sekundaari adjektiivi. (Adjektiivien transpositiota Erelt on käsitelty artikkelissaan »Some Notes on the Classification of Words and the Transposition of Adjectives in Estonian», joka on ilmestynyt kirjassa *Grammar and Semantics*, Tallinn 1980.)

Toisessa tyyppissä on attribuuttina verbaalinen adjektiivi: *v ~ tav* -partisiippi, esim. *raamatul lugev poiss, kirjutataw kiri*. Voidaan katsoa, että näissä esiintyy sekundaari adjektiivi, joka on verbin täydellisen syntaktisen adjektiivistumisen tulos: partisiippi kongruoi pääsanana kanssa kuten adjektiivi; on kyllä syntaktisia ja semanttisia seikkoja, jotka viittaavat partisiippien verbiyteen. Tämän tyyppin proposition rakenteessa predikaatti ilmaisee prosessia, argu-

menti substanssia; tällöisiä rakenteita onkin vanhastaan kuvattu seuraavasti: *jooksev poiss — poiss, kes jookseb. v ~ tav* -partisiippiä ei tavata predikatiivina paitsi esim. sellaisessa tapauksessa kuin *käekiri on loetav*, jossa partisiippi ei ilmaise prosessia vaan (semanttisesti adjektiivistuneena) ominaisuutta. Vastaava kaksimerkityksinen attribuuttirakenne *loetav käekiri* vietään merkityksessä 'lukemisen arvoinen käsikirja' perustyyppiin.

Seuraavaa tyyppiä edustavat esim. *karusnahkne kasukas, meditsiiniline uurimus, teaduslik sekretär, sünsed seadused, eilne ajaleht*, joiden attribuuttina on syntaktisesti adjektiivistunut substantiivi tai myös adverbi. Tämöisten tapausten proposition rakenteessa ovat argumentteina pääsana ja se substantiivi, jonka johdos adjektiivi on. Predikaatti puolee rakenteen merkityksessä, ja usein sitä on vaikea tarkoin määrittää menemättä syvälle semantiikkaan. Esim. ensimmäisen esimerkin proposition rakenne on samantapainen kuin lauseen *Kasukas on valmistatud karusnahast*. Tämän rakenteen kuten edellisenkin tyyppin attribuutit ovat suhdeadjektiiveja, joille on ominaista restrktiivinen (klassifioiva) funktio: ne eivät yleensä esiinny predikatiivina, esim. **See sekretär on teaduslik*. Toisaalta voi sanoa *See kasukas on karusnahkne, See uurimus on meditsiiniline, See ajaleht on eilne*. Ensimmäinen tapaus ei tekijän mukaan ole varsinainen poikkeus, koska siinä voi olla kyse adjektiivin (sekundaarista) verbaalistumisesta; toinen tapaus taas lienee oikeastaan attribuuttirakenne: adjektiivin jäljessä voidaan ajatella toistettavaksi sana *uurimus* ilman, että lause tuntuisi kovin epäkorrektilta (toisin kuin esim. *?See poiss on rumal poiss*).

Myös viimeinen attribuuttirakennetyypipi on problemaattinen. Siinä on attribuuttina määränadjektiivi, joka ilmaisee prosessin ominaisuutta; proposition aspektista esim. rakenne *hea laulja* vastaa lausetta *keegi, kes laulab hästi*. Tässä siis katsotaan, että attribuutti on adverbien syntaktisen adjektiivistumisen tulos. Esim. tapauksissa *võimalik eksimus, usin töö* attribuutti määrittää absolutteisesti pääsanana ilmaisemaa prosessia, mutta sellaisissa tapauksissa, joissa on mukana semanttinen subjekti, on kyse epäsuorasta attribuutiosta, jossa attribuut-

Kirjallisuutta

ti määrittää henkilöä tämän toiminnan kautta. Epäsuora attribuutio on ominaista vain attribuuttirakenteelle. Predikatiivilauseessa viitataan konkreettiseen objektiin, joten (puhdas) määränadjektiivi ei ole siinä käypä: *täielik idioot* — *(See) *idioot on täielik*. *Vana merimees* on kaksimerkityksinen adjektiivin semantiikan vuoksi: *vana* voi ilmaista sekä määrää että ominaisuutta; tulkinnan mukaan esimerkki kuuluu joko tähän kohtaan tai perustyyppiin. Kiistatta lukeutuvat tekijänkin mukaan tämän kohdan tyyppiin *ja*-johtimisen pääsanana sisältävät rakenteet, samoin sellaiset, joiden pääsanana on merkitykseltään vähän erikoistuneempi johdos, esim. *suur joodik* ja ehkä vielä *hea insener*, mutta suuria vaikeuksia aiheuttaa sellaisten rakenteiden luokittaminen, joiden prosessiainesta on hankalampi määritellä, esim. *hea isa*. Luullakseni tässä aletaan olla perustyyppin rajoilla. Selvät suoran attribuution tapaukset sekä sellaiset rakenteet, joiden merkitys sisältää instrumentin, tekijä lukee perustyyppiin: niiden rinnalla esiintyy aina predikatiivilause, esim. instrumenttitapaus on *halb kirjutusmasin*.

Mati Ereltin tutkielmassa saavat huomattavaa lisävalaistusta kielioppien ja tutkimusten laiminlyömät attribuuttirakenteiden sekundaarityypit ja niiden suhde perustyyppiin ja predikatiivilauseeseen. Propositionaliansalysissä attribuuttirakenteet ryhmittyvät selkeästi neljäksi tyyppiiksi. Osoittautuu kuitenkin, että propositionaliansalysissä ei ole aina riittävä tyyppityskriteeri: eräillä sekundaareilla rakenteilla on perustyyppin kanssa yhteisenä ominaisuutena se, että niiden rinnalle on mahdollista muodostaa predikatiivilause. Tämän ilmiön selitystä on varmaankin etsittävä mm. syvempää semantiikasta.

JARMO ELOMAA